Package ‘loadings’

September 2, 2022

Type Package

Title Loadings for Principal Component Analysis and Partial Least Squares

Version 0.2.1

Author Hiroyuki Yamamoto

Maintainer Hiroyuki Yamamoto <h.yama2396@gmail.com>


License LGPL-3

Encoding UTF-8

LazyData true

Imports geigen

NeedsCompilation no

Repository CRAN

Date/Publication 2022-09-02 07:40:02 UTC

R topics documented:

covid19 .................................................. 2
fasting .................................................. 3
greentea ............................................... 3
multiplsrog_loading .................................. 4
multipls_geigen ...................................... 5
multipls_loading ..................................... 6
multipls_rog ......................................... 8
ospca_loading ....................................... 9
os_pca ............................................... 10
pca_loading .......................................... 12
plsrog_loading ...................................... 13
pls_loading .......................................... 14
covid19

Proteome and metabolome data of COVID-19 serum samples.

Description

This is proteome and metabolome data of COVID-19 serum samples. This dataset includes three groups of healthy control, mild and severe COVID-19.

Usage

data(covid19)

Format

The list object "covid19" contains the following elements:

X : List of data matrix that include samples in each rows and protein or metabolite in each columns for proteomics and metabolomics.

Y : Dummy matrix that include group information 0,1 in each columns.

D : Example of differential matrix used in PLS-ROG or multiset PLS-ROG.

tau : Example of strength parameter matrix used in mutliset PLS or multiset PLS-ROG. All weights were set as the same value of 0.1.

References


Examples

data(covid19)
Description

This data includes metabolome data of liver samples from normal and 12 hour fasted mice measured by capillary electrophoresis time-of-flight mass spectrometry. The data matrix contains 10 samples (5 samples in each group) and 282 metabolites.

Usage

data(fasting)

Format

The list object "fasting" contains the following elements:
X : Data matrix that include metabolites in each columns.
Y : Dummy matrix that include group information 0,1 in each columns.

Source


Description

This includes metabolome data of Japanese green tea measured by gas chromatography mass spectrometry. The data matrix contains 30 samples and 225 metabolites.

Usage

data(greentea)

Format

The list object "greentea" contains the following elements:
X : Data matrix that include metabolites in each columns.
Y : Dummy matrix that include group information 0,1 in each columns.
D : Differential matrix (Second order).
M : Averaging matrix for repeated data.
y : Ranking of taste in competitive exhibition.
**multiplsrog_loading**

**Description**

This function computes multiset partial least squares rank order of groups (PLS-ROG) loading from the result of the "multipls_rog" function.

**Usage**

```r
multiplsrog_loading(multiplsrog)
```

**Arguments**

- `multiplsrog` The following variables (P,T,Q and U) are included in the multiplsrog object.
  - `P`: A list of matrix with Multiset PLS-ROG coefficients for the explanatory variables in each column for each dataset
  - `T`: A list of matrix with Multiset PLS-ROG scores for the explanatory variables in each column for each dataset
  - `Q`: A matrix with Multiset PLS-ROG coefficients for the response variable in each column
  - `U`: A matrix with Multiset PLS-ROG scores for the response variable in each column
  - `tau`: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information.

**Details**

The multiset PLS-ROG loading and the p-value by statistical hypothesis testing is added to the multiplsrog object returned by the "multipls_rog" function in this package.

**Value**

The return value is a list object that contains the following elements:

- `R`: A list of matrix with Multiset PLS-ROG loadings in each column for each dataset
- `p.value`: A list of matrix with p-value of multiset PLS-ROG loading in each column for each dataset

**Author(s)**

Hiroyuki Yamamoto
References


Examples

```r
data(whhl)
X <- whhl$X
Y <- whhl$Y
D <- whhl$D
tau <- whhl$tau

multiplsrog <- multipls_rog(X,Y,tau,D)
# multiplsrog <- multipls_rog(X,Y,tau,D, kappa=0.999)

multiplsrog <- multiplsrog_loading(multiplsrog)
multiplsrog$loading$R
multiplsrog$loading$p.value
```

---

**multipls_geigen**

*Multiset PLS: Multiset partial least squares*

**Description**

This function performs Multiset partial least squares (Multiset PLS). In this function, data matrix is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

**Usage**

```r
multipls_geigen(X, Y, tau)
```

**Arguments**

- **X**: List of data matrix that include variables in each columns.
- **Y**: Dummy matrix that include group information 0,1 in each columns.
- **tau**: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information.

**Details**

Diagonal elements of matrix tau must be 0.
**Value**

The return value is a list object that contains the following elements:

- **P**: A list of matrix with Multiset PLS coefficients for the explanatory variables in each column for each dataset
- **T**: A list of matrix with Multiset PLS scores for the explanatory variables in each column for each dataset
- **Q**: A matrix with Multiset PLS coefficients for the response variable in each column
- **U**: A matrix with Multiset PLS scores for the response variable in each column
- **tau**: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information (same as input value).

**Author(s)**

Hiroyuki Yamamoto

**References**


**Examples**

```r
data(whhl)
X <- whhl$X
Y <- whhl$Y
tau <- whhl$tau

multipls <- multipls_geigen(X,Y,tau)
```

**Description**

This function computes multiset partial least squares (PLS) loading from the result of the "multipls_geigen" function.

**Usage**

`multipls_loading(multipls)"
Arguments

multipls: The following variables (P,T,Q and U) are included in the multipls object.
P: A list of matrix with Multiset PLS coefficients for the explanatory variables in each column for each dataset
T: A list of matrix with Multiset PLS scores for the explanatory variables in each column for each dataset
Q: A matrix with Multiset PLS coefficients for the response variable in each column
U: A matrix with Multiset PLS scores for the response variable in each column
tau: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information.

Details

The multiset PLS loading and the p-value by statistical hypothesis testing is added to the multipls object returned by the "multipls_geigen" function in this package.

Value

The return value is a list object that contains the following elements:
R: A list of matrix with Multiset PLS loadings in each column for each dataset
p.value: A list of matrix with p-value of multiset PLS loading in each column for each dataset

Author(s)

Hiroyuki Yamamoto

References


Examples

data(whhl)
X <- whhl$X
Y <- whhl$Y
tau <- whhl$tau

multipls <- multipls_geigen(X,Y,tau)

multipls <- multipls_loading(multipls)
multipls$loading$R
multipls$loading$p.value
Description

This function performs multiset partial least squares with rank order of groups (Multiset PLS-ROG). In this function, data matrix is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

```
multipls_rog(X,Y,tau,D,kappa)
```

Arguments

- **X**: List of data matrix that include variables in each columns.
- **Y**: Dummy matrix that include group information 0,1 in each columns.
- **tau**: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information.
- **D**: Differential matrix.
- **kappa**: The smoothing parameter (default : 0.999).

Details

Diagonal elements of matrix tau must be 0.

Value

The return value is a list object that contains the following elements:

- **P**: A list of matrix with Multiset PLS-ROG coefficients for the explanatory variables in each column for each dataset
- **T**: A list of matrix with Multiset PLS-ROG scores for the explanatory variables in each column for each dataset
- **Q**: A matrix with Multiset PLS-ROG coefficients for the response variable in each column
- **U**: A matrix with Multiset PLS-ROG scores for the response variable in each column
- **tau**: Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information (same as input value).

Author(s)

Hiroyuki Yamamoto
ospca_loading

References

Examples
data(whhl)
X <- whhl$X
Y <- whhl$Y
D <- whhl$D
tau <- whhl$tau

multiplsrog <- multipls_rog(X,Y,tau,D)
# multiplsrog <- multipls_rog(X,Y,tau,D, kappa=0.999)

ospca_loading

Orthogonal smoothed principal component loading

Description
This function computes orthogonals smoothed principal component (OS-PC) loading from the result of the "os_pca" function. This is also wrapper function of pls_loading function.

Usage
ospca_loading(ospca)

Arguments
ospca
The following variables (P,T,Q and U) are included in the ospca object.
P : A matrix with OS-PC loading in each column
T : A matrix with OS-PC score in each column
MT : A matrix with averaging OS-PC score for repeated data in each column (If not for repeated data, the matrix MT equals to the matrix T)
Q : A matrix with OS-PC loading for auxiliary variable in each column
U : A matrix with OS-PC score for auxiliary variable in each column

Details
The OS-PC loading and the p-value by statistical hypothesis testing is added to the ospca object returned by the "os-pca" function.

Value
The return value is a list object that contains the following elements:
R : OS-PCA loading (Correlation coefficient between OS-PC score of auxiliary variable and each variables in data matrix.)
p.value : p-value of OS-PCA loading
os_pca

Orthogonal smoothed principal component analysis

Description

This function performs orthogonal smoothed principal component analysis (OS-PCA). In this function, data matrix is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

```
os_pca(X,D,kappa,M)
```
Arguments

- **X**: Data matrix that include variables in each columns.
- **D**: Differential matrix.
- **kappa**: The smoothing parameter (default: 0.999).
- **M**: Averaging matrix for repeated data (default: Identity matrix).

Details

The kappa represents the degree of smoothing. The value of kappa increases, the strength of the smoothing increases.

Value

The return value is a list object that contains the following elements:
- **P**: A matrix with OS-PC loading in each column
- **T**: A matrix with OS-PC score in each column
- **MT**: A matrix with averaging OS-PC score for repeated data in each column (If not for repeated data, the matrix MT equals to the matrix T)
- **Q**: A matrix with OS-PC loading for auxiliary variable in each column
- **U**: A matrix with OS-PC score for auxiliary variable in each column

Author(s)

Hiroyuki Yamamoto

References


Examples

```r
# metabolic turnover data
data(turnover)
X <- turnover$X
D <- turnover$D
ospca <- os_pca(X,D,0.999)

# metabolome data
data(greentea)
X <- greentea$X
D <- greentea$D
M <- greentea$M
ospca <- os_pca(X,D,0.1,M)
```
Description

This function computes principal component (PC) loading from the result of the "prcomp" function. (The "princomp" function is not supported. For "princomp" function, the "loadings" function in stats package should be used.) In this function, data matrix is should be scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

pca_loading(pca)

Arguments

pca

The return object of the "prcomp" function.

Details

The PC loading and the p-value by statistical hypothesis testing is added to the pca object returned by the "prcomp" function.

Value

The return value is a list object that contains the following elements:

R : PC loading (Correlation coefficient between PC score and each variables in data matrix.)

p.value : p-value of PC loading

Author(s)

Hiroyuki Yamamoto

References


Examples

data(fasting)
X <- fasting$X

pca <- prcomp(X, scale=TRUE)
pca <- pca_loading(pca)
pca$loading$R # PC loading
pca$loading$p.value # p-value
Function: plsrog_loading

### Description

This function computes partial least squares rank order of groups (PLS-ROG) loading from the result of the "pls_rog" function. PLS-ROG loading can also be computed from the result of the "pls_rog" function.

### Usage

```r
plsrog_loading(plsrog)
```

### Arguments

- **plsrog**: The following variables (P,T,Q and U) are included in the plsrog object.
  - P: A matrix with PLS-ROG loading for explanatory variable in each column
  - T: A matrix with PLS-ROG score for explanatory variable in each column
  - Q: A matrix with PLS-ROG loading for response variable in each column
  - U: A matrix with PLS-ROG score for response variable in each column

### Details

The PLS-ROG loading and the p-value by statistical hypothesis testing is added to the pls object returned by the "pls_rog" function in this package.

### Value

The return value is a list object that contains the following elements:

- **R**: PLS-ROG loading (Correlation coefficient between PL-ROG score of response variable and each variables in data matrix.)
- **p.value**: p-value of PLS-ROG loading

### Author(s)

Hiroyuki Yamamoto

### References

Examples

data(whhl)
X <- whhl$X$liver
Y <- whhl$Y
D <- whhl$D

plsrog <- plsrog(X,Y,D)

plsrog <- plsrog_loading(plsrog)
plsrog$loading$R
plsrog$loading$p.value

---

**pls_loading**

Partial least squares loading

---

**Description**

This function computes partial least squares (PLS) loading from the result of the "pls_svd" function. PLS loading can also be computed from the result of the "pls_eigen" function in chemometrics package.

**Usage**

`pls_loading(pls)`

**Arguments**

- **pls** The following variables (P,T,Q and U) are included in the pls object.
  
  - **P** : A matrix with PLS loading for explanatory variable in each column
  
  - **T** : A matrix with PLS score for explanatory variable in each column
  
  - **Q** : A matrix with PLS loading for response variable in each column
  
  - **U** : A matrix with PLS score for response variable in each column

**Details**

The PLS loading and the p-value by statistical hypothesis testing is added to the pls object returned by the "pls_svd" function in this package or "pls_eigen" function in chemometrics package.

**Value**

The return value is a list object that contains the following elements:

- **R** : PLS loading (Correlation coefficient between PLS score of response variable and each variables in data matrix.)
- **p.value** : p-value of PLS loading
Author(s)

Hiroyuki Yamamoto

References


Examples

data(whhl)
X <- whhl$X$liver
Y <- whhl$Y

pls <- pls_svd(X,Y)

pls <- pls_loading(pls)
pls$loading$R
pls$loading$p.value

Description

This function performs partial least squares with rank order of groups (PLS-ROG). In this function, data matrix is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

pls_rog(X,Y,D,kappa)

Arguments

X   Data matrix of explanatory variables that include variables in each columns.
Y   Dummy matrix that include group information 0,1 in each columns.
D   Differential matrix.
kappa  The smoothing parameter (default : kappa = 0.999).

Details

The kappa represents the degree of smoothing. The value of kappa increases, the strength of the smoothing increases.
Value

The return value is a list object that contains the following elements:
P : A matrix with PLS loading for explanatory variable in each column
T : A matrix with PLS score for explanatory variable in each column
Q : A matrix with PLS loading for response variable in each column
U : A matrix with PLS score for response variable in each column

Author(s)

Hiroyuki Yamamoto

References


Examples

data(whhl)
X <- whhl$X$liver
Y <- whhl$Y
D <- whhl$D

plsrog <- pls_rog(X,Y,D)

---

pls_svd                     Partial least squares

Description

This function performs partial least squares. In this function, data matrix for explanatory variable is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

pls_svd(X,Y)

Arguments

X   Data matrix of explanatory variables that include variables in each columns.
Y   Dummy matrix that include group information 0,1 in each columns.

Details

This function is wrapper function of "pls_rog" function that the smoothing parameter kappa=0.
Value

The return value is a list object that contains the following elements:
P : A matrix with PLS loading for explanatory variable in each column
T : A matrix with PLS score for explanatory variable in each column
Q : A matrix with PLS loading for response variable in each column
U : A matrix with PLS score for response variable in each column

Author(s)

Hiroyuki Yamamoto

References


Examples

data(whhl)
X <- whhl$X$liver
Y <- whhl$Y
pls <- pls_svd(X,Y)

<table>
<thead>
<tr>
<th>turnover</th>
<th>Metabolomic Turnover analysis of the yeast culture mediums</th>
</tr>
</thead>
</table>

Description

This includes differential isotopomer ratio from metabolomic turnover analysis of yeast cultures by gas chromatography mass spectrometry. The data matrix contains 11 sampling time for three groups of different yeast strains and culture condition (33 samples total) and 60 peaks.

Usage

data(turnover)

Format

The list object "turnover" contains the following elements:
X : Data matrix of explanatory variables that include variables in each columns.
Y : Dummy matrix that include group information 0,1 in each columns.
D : Differential matrix (Second order).
References

Examples
data(turnover)

unsvmultipls_loading  Unsupervised multiset partial least squares loading

Description
This function computes unsupervised multiset partial least squares loading from the result of the "unsv_multipls" function.

Usage
unsvmultipls_loading(unsvmultipls)

Arguments
unsvmultipls  The following variables (P,T) are included in the unsvmultipls object.
P : A list of matrix with Multiset PLS coefficients in each column for each dataset
T : A list of matrix with Multiset PLS scores in each column for each dataset

Details
The unsupervised multiset PLS loading and the p-value by statistical hypothesis testing is added to the unsvmultipls object returned by the "unsv_multipls" function in this package.

Value
The return value is a list object that contains the following elements:
R : A list of matrix with Unsupervised multiset PLS loadings in each column for each dataset
p.value : A list of matrix with p-value of Unsupervised multiset PLS loading in each column for each dataset

Author(s)
Hiroyuki Yamamoto

References
unsv_multipls

Examples

data(whhl)
  X <- whhl$X
  tau <- rbind(c(0, 1/4, 1/4, 1/4), c(1/4, 0, 1/4, 1/4), c(1/4, 1/4, 0, 1/4), c(1/4, 1/4, 1/4, 0))

  unsvmpls <- unsv_multipls(X, tau)
  unsvmultipls <- unsvmultipls_loading(unsvmpls)
  unsvmultipls$loading$R
  unsvmultipls$loading$p.value

unsv_multipls  Unsupervised multiset partial least squares

Description

This function performs Unsupervised multiset partial least squares. In this function, data matrix is automatically scaled to zero mean and unit variance (i.e. autoscaling) for each variables.

Usage

unsv_multipls(X, tau)

Arguments

X       List of data matrix that include variables in each columns.
tau     Matrix for strength parameter of the connection between omics datasets or between omics dataset and group information.

Details

Diagonal elements of matrix tau must be 0.

Value

The return value is a list object that contains the following elements:
P : A list of matrix with Multiset PLS coefficients in each column for each dataset
T : A list of matrix with Multiset PLS scores in each column for each dataset

Author(s)

Hiroyuki Yamamoto

References


**Examples**

```r
data(whhl)
X <- whhl$X
tau <- rbind(c(0, 1/4, 1/4, 1/4), c(1/4, 0, 1/4, 1/4), c(1/4, 1/4, 0, 1/4), c(1/4, 1/4, 1/4, 0))
unsvmpls <- unsv_multipls(X, tau)
```

---

**whhl**

*Metabolome analysis of plasma, liver, heart and brain samples from rabbits with hyperlipidemia.*

**Description**

This is metabolome data from plasma, liver, heart and brain samples of normal, hyperlipidemic, and statin-treated rabbits measured by capillary electrophoresis time-of-flight mass spectrometry. The data matrix contains 9 samples (3 samples in each groups).

**Usage**

```r
data(whhl)
```

**Format**

The list object "whhl" contains the following elements:

- **X**: List of data matrix that include samples in each rows and metabolites in each columns for plasma, liver, heart and brain samples.
- **Y**: Dummy matrix that include group information 0, 1 in each columns.
- **D**: Example of differential matrix used in PLS-ROG or multiset PLS-ROG.
- **tau**: Example of strength parameter matrix used in mutliset PLS or multiset PLS-ROG. All weights were set as the same value of 0.1.

**References**


**Examples**

```r
data(whhl)
liver <- whhl$X$liver
```
Index

* datasets
  covid19, 2
  fasting, 3
  greentea, 3
  turnover, 17
  whhl, 20

* multivariate
  multipls_geigen, 5
  multipls_loading, 6
  multipls_rog, 8
  multiplsrog_loading, 4
  os_pca, 10
  ospca_loading, 9
  pca_loading, 12
  pls_loading, 14
  pls_rog, 15
  pls_svd, 16
  plsrog_loading, 13
  unsv_multipls, 19
  unsvmultipls_loading, 18

  covid19, 2
  fasting, 3
  greentea, 3
  multipls_geigen, 5
  multipls_loading, 6
  multipls_rog, 8
  multiplsrog_loading, 4
  os_pca, 10
  ospca_loading, 9
  pca_loading, 12
  pls_loading, 14
  pls_rog, 15
  pls_svd, 16
  plsrog_loading, 13
  unsv_multipls, 19
  unsvmultipls_loading, 18

  turnover, 17
  unsv_multipls, 19
  unsvmultipls_loading, 18
  whhl, 20